三维片上网络体系结构研究综述

李晨马胜王璐郭阳

摘 要 伴随着三维集成电路的迅速发展,三维片上网络受到国内外研究者的广泛关注.三维片上网络主要用于 实现三维堆叠芯片的互连通信,为三维集成电路提供超低的延迟和竖直方向上超高的带宽,从而解决系统集成度 增加导致的通信瓶颈问题,有利于克服存储墙问题并提高三维堆叠芯片的性能.文章介绍了三维集成电路研究现 状及其结构优势,分析三维片上网络体系结构的特点和存在的问题,包括竖直方向的单跳传播问题、路由器交叉开 关的复杂度控制问题以及热效应的控制问题,从系统层、微结构层和电路层对三维片上网络体系结构的研究热点 及其实例进行了深入分析,最后对三维片上网络所面临的挑战和设计方法进行了总结和展望.

关键词 计算机体系结构;三维堆叠;三维集成电路;三维片上网络中图法分类号 TP393 DOI号 10.11897/SP.J.1016.2016.01812

A Survey on Architecture for Three-Dimensional Networks-on-Chip

LI Chen MA Sheng WANG Lu GUO Yang

(College of Computer, National University of Defense Technology, Changsha 410073)

Abstract 3D integrated circuit, stacked multiple silicon layers vertically, can significantly reduce latency and energy consumption. With the development of the 3D integrated circuit, 3D Networks-on-Chip (NoCs) has widely received attention from researchers all over the world. The 3D NoC is used for interconnection among 3D stacking chips. It provides ultra low latency and ultra high bandwidth for the 3D integrated circuit through vertical direction. It can address the bottleneck of communication caused by the increasing of system integration, then the problem of memory wall can also be addressed. It will largely improve the performance of 3D stacked chips. This article first reviews the current research progress and advantages in 3D structure, analyzes features and challenges of 3D NoC, including single-hop transferring problem through vertical direction, the complexity control of router crossbar and the heat control problem, and then introduces research hotspots of 3D NoC from the system layer, microarchitecture layer and circuit layer, including the introduction and analyses of examples. These research hotspots are extended from 3 basic NoC portions, including topology, routing algorithm and router microarchitecture. Finally, some open issues and future directions in 3D NoC are discussed.

Keywords computer architecture; 3D stacking; 3D integrated circuit; 3D NoC

1 引 言

摩尔定律实际上是半导体集成电路行业的一条

学习曲线,一直引导和促进集成电路工艺技术的发展和进步.在2004年以前,摩尔定律主要通过工艺 缩减来维持,然而当工艺缩减到20nm后,成本下降 的空间非常有限,开发难度也陡然增加,晶体管尺寸

收稿日期:2015-05-20;在线出版日期:2015-10-10.本课题得到国家自然科学基金(61133007,61303065)、教育部博士点基金(20134307120028) 资助. 李 晨,男,1989年生,博士研究生,主要研究方向为片上网络. E-mail: lichen@nudt. edu. cn. 马 胜(通信作者),男,1986年生,博士,助理研究员,中国计算机学会(CCF)会员,主要研究方向为计算机体系结构. E-mail: masheng@nudt. edu. cn. 王 璐,女,1991年生,硕士研究生,主要研究方向为计算机体系结构. 郭 阳,男,1971年生,博士,研究员,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为微处理器设计验证技术.

的线性缩减有所减缓^[1],因此出现了多核、虚拟化等 并行策略来维持这一学习曲线.随着技术的继续发展,多核、多线程和虚拟化等策略又会遇到存储墙的 问题^[2],需要提供更高的存储带宽和更大的存储空 间来保证其性能.在这个背景下,急需找到一种新的 方式来维持这一学习曲线的速率,因此,三维芯片堆 叠的方法应运而生^[3].

三维集成技术是一种新的集成电路工艺,其将 多个硅片垂直堆叠并以三维封装的方式封装成一个 芯片[4],是在工艺尺寸缩减受限的情况下提高系统 性能的一种新的方式[5],从延伸摩尔定律和超越摩 尔定律两方面实现芯片上晶体管密度和芯片性能的 大幅度提升[6-8]. 三维集成电路具有以下 4 点优势: (1) 更短的互连线长, 三维集成电路技术最大的吸 引力就在于工艺缩减受限的情况下,通过芯片堆 叠,降低互连线的平均长度[9],从而缩短延迟并降低 功耗;(2)更高的存储带宽. 三维集成电路通过对存 储部件的堆叠,一方面提高了 Cache 容量,另一方面 减少了平均访存时间,从而提升了吞吐率和能效; (3)支持异质集成的实现. 三维集成电路增加的一 维空间使得芯片可以在更大的密度下实现更高的性 能,也使集成电路的异质集成能力得以更好地体现, 即体系结构设计者可以通过灵活的方法实现各种异 质结构的堆叠来满足功能的多样性;(4)更小的系 统尺寸. 三维集成电路技术将二维芯片竖直堆叠, 通过纵向垂直互连进行通信.该方式在提升了芯片 集成度的同时缩小了芯片面积.

三维片上网络通过三维立体的互连通信,为三 维集成电路提供超低的延迟和竖直方向上超高的带 宽,在现阶段对于提高三维芯片性能并控制热效应 问题有着非常重要的意义.三维集成电路必将是未 来集成电路的发展方向,三维片上网络作为三维集 成电路最重要的互连通信方式,具有非常广阔的发 展前景.

基于此,本文针对目前三维片上网络的研究热 点,从系统层、微结构层和电路层三个层次对三维片 上网络体系结构进行热点综述.分析了三维拓扑结 构和路由器微结构的发展轨迹;对路由算法、垂直互 连的 TSV 优化设计以及热管理等方面研究进行了 分类总结;同时还对非常有前景的 2.5 维片上互连 硅中介技术进行介绍和讨论;此外,对于三维片上网 络在设计方法学、缓存分配以及新型堆叠方式等方 面的研究也作了分析介绍.通过这些文献综述和讨 论,分析现有研究的优势和不足,总结三维片上网络 研究所面临的问题和挑战,并针对这些问题和挑战 提出下一步的研究前景和发展方向.

2 三维片上网络体系结构特点与 研究层次

三维片上网络体系结构的最大特点是三维集成 堆叠以及与此相适应的竖直通信链路.目前,三维集 成逻辑堆叠的商用芯片产品并未普及,主要是以存 储堆叠为主,但按可能批量生产的时间顺序可以分 为存储堆叠、硅中介层、存储层堆叠逻辑层和逻辑三 维堆叠4种形式[10],这些三维堆叠都是通过纵向的 垂直互连进行通信.纵向垂直互连技术包括引线接 合法(wire bonding)、微焊点(microbump)、无触点 (contactless)以及硅通孔连接(TSV)等几种方式, 这几种连接方式在纵向连接密度和实现特点方面各 具特色^[11]. 就目前来看,采用 TSV 的垂直互连技术 通过在硅晶圆上穿孔并以钨金属填注进行芯片间互 连,相对来说是最有前景的垂直互连技术,最有潜力 提供最大的互连密度,虽然其制造成本目前还比较 高,但由于其出色的性能,TSV 垂直互连技术有望 获得最广泛的应用[12].

片上网络作为一种在芯片内部通过拓扑结构将 各个模块连接,并采用路由和分组交换技术进行数 据交换的通信方式,可有效解决复杂 SoC 通信问 题[13-14]. 三维片上网络是指将资源节点分布在竖直 堆叠在一起的不同的物理层上,并采用了三维立体 的结构实现节点间的互连通信,实现低延迟、高带宽 的片上网络结构[15-16]. 三维片上网络的体系结构设 计对于网络延迟、吞吐率、功耗、应用映射以及芯片 的面积、复杂度都有着非常重要的影响[17-18]. 三维片 上网络和二维片上网络最大的区别在于新增加的连 接不同层的竖直链路,纵向竖直链路可以实现传统 二维片上网络的平面结构所无法实现的性能,因此, 如何利用好这些特别的竖直链路成为三维片上网络 的研究重点.这就要求设计者选择一个合适的竖直 链路体系结构,并设计高性能低开销的路由器来减 少竖直方向上的节点跳数.

三维片上网络最关键的研究目标就是如何在相 对较小的面积和功耗开销下,为三维芯片提供更高 的互连性能.相对于二维片上网络,三维片上网络系 统结构层的大部分研究集中在解决3个问题:

(1)竖直方向的单跳传播问题,即在层间互连 中,由于竖直方向距离相对于水平方向可忽略不计, 因此要求任意两层之间在竖直方向都采用单跳传播 以降低延迟.

(2)路由器交叉开关的复杂度控制问题,由于 三维结构增加一维需要增加交叉开关的端口数量, 进而大幅增大交叉开关的复杂度,使得硬件开销不 可接受,因此需要对交叉开关进行简化设计以降低 硬件开销.

(3) 热效应的控制问题, 三维堆叠技术使得集成电路晶体管密度增大, 单位面积产生的热量也相应增大, 同时三维堆叠增大了散热难度, 所以热效应成为必须控制的问题.

上述问题的研究产生了几大研究热点,这些研究热点分别分布在系统层、微结构层和电路层,概括如图1所示.这种分层的视角可以更好地理解三维片上网络研究层次.在系统层,主要是对片上网络进行全局性研究,如拓扑结构、路由算法、硅中介技术^[19]、



图 1 三维片上网络研究层次图

设计方法学、新型芯片堆叠方式以及热管理的部分 内容;在微结构层研究方面,主要包括路由器缓存分 配、路由器微体系结构研究以及热管理部分涉及路 由器结构的设计研究;在电路层方面,主要是以垂直 TSV 互连相关内容的研究为主.下面将针对这些体 系结构研究热点进行分析.

3 三维片上网络体系结构研究进展

3.1 拓扑结构

片上网络主要由链路通道和路由器节点组成, 而网络拓扑结构则决定了其物理布局和路由器节点 与链路的连接方式. 拓扑结构的评价指标有节点的 度、平均节点跳数、最大通道负载以及路径多样性 等.对于三维片上网络来说,增加的一维纵向竖直连 接不同于传统二维片上网络的平面连接方式,为了 利用好纵向连接的短距离和高带宽的特性,垂直通 信在三维拓扑结构的设计上也需要特别考虑,目前 对三维片上网络拓扑结构的研究非常多[26,28-30],其 发展路线如图 2 所示,主要可以分为两类,一类将 二维片上网络已有的拓扑结构直接应用于三维片 上网络,发挥三维立体结构的空间优势,实现资源 利用的最大化;另一类则是根据三维片上网络的 特点单独设计,包括利用竖直方向相对于二维平 面上可忽略不计的距离、竖直方向上 TSV 通信的 高带宽等.下面将介绍这两类拓扑结构设计思路的 研究现状.



图 2 三维拓扑结构发展路线

三维对称网络结构^[20]是最基本的三维网络结构,它由二维网络结构向竖直方向直接拓展而成.如 图 3(a)所示,该结构相当于在二维网络结构的路由 器上增加两个端口,分别是在竖直方向朝上和朝下 的两个端口.显然,增加这些端口需要相应地增加缓存、仲裁器以及交叉开关的规模.所谓对称的网络结构是指报文在片间和片内都是每节点一跳地传输,即报文从一个四层芯片的底层某节点传输到顶层某





(c) 三维纤毛片上网络结构

图 3 三维拓扑结构

节点至少需要经过3跳.三维对称网络结构非常简 单、易实现,但存在两个主要缺陷:一是片间多跳传 输浪费了其相对于片内可忽略不计的距离;另一个 是增加额外的端口,需要一个7×7的交叉开关,开 销非常大.

三维对称网络在竖直方向上按跳传输是极大的 资源浪费,因此,文献[21]提出了三维片上网络-总 线混合结构以解决该问题.如图 3(b)该结构在片内 采用包交换网络,片间采用总线结构,通过这种混合 机制,很好地解决了片间多跳传输的问题.片内的包 交换网络与二维片上网络相似,而在片间采用的总 线则是被称为"通信柱"的动态时分多路访问总线 (dTDMA).这些通信柱为任意两层芯片的节点提 供单跳通信.限制通信柱数量的因素主要包括通孔 的密度和路由器复杂度.采用三维片上网络-总线混 合结构相对于三维对称片上网络结构解决了片间多 跳传输的问题,降低了硬件复杂度,将交叉开关规模 从7×7降低为6×6.然而,该结构在竖直方向上不 支持并行通信,在高流量下冲突严重,因此对于节点 数目和层数的限制非常严格.

三维片上网络-总线混合结构在竖直方向存在 不支持并行通信的问题,而三维纤毛片上网络结构^[22]很好地解决了这个问题.它是一种单 Switch 连接多 IP 模块的三维片上网络结构.如图 3(c)所 示,该 3×3×3 的三维纤毛结构由 9 个 Switch 和 27 个 IP 模块组成,其中每个 Switch 连接 3 个 IP 模 块. Switch 只存在于三维结构的最底层,各层的 IP 模块都与最底层对应的 Switch 相连接. 在层内互连 中,只有接近散热模块的底层具有较好的连通性,其 他层的层内通信都需要经过底层 Switch 来完成. 这 种结构面积和功耗开销低,在一些特定的流量模型 下能发挥较好的性能,但单 Switch 多 IP 模块的设 计降低了该结构的连通性,因此全局带宽是该结构 的瓶颈,影响其整体性能.

XNOTS 是一种通过交叉开关 Switch 进行层 间互连的片上网络结构^[23],各层内可采用任意直接 拓扑逻辑结构,而层间则通过交叉开关 Switch 非直 接相连.如图 4 所示,这种交叉开关 Switch 被称为 路由柱.XNOTS 三维片上网络结构的另一特点就 是除了底层外,无需保证各层内部任意两节点之间 的连通性,这类似于三维纤毛结构,各层之间可根据 性能和开销需要独立定制设计.由于散热器离底层



图 4 XNOTS 路由柱

较近,易于对底层散热,因此,对于不连通的上层网络节点,均通过底层网络连接.这种设计还采用了热感知路由算法,具有较好的散热性能和较高的吞吐率,但是竖直方向上的交叉开关设计对于面积和功耗的需求相对较高.

为了降低三维片上网络的报文传播延迟,Xu等 人^[24]提出了一种如图 5(a)所示的低开销路由器组 成的低直径三维片上网络设计.在采用与传统三维 Mesh 片上网络相同的 7 阶路由器的基础上,通过 在 Cache 层使用长连线的方式组成网络直径非常 低的三维片上网络结构.由于计算核所在层的内 部 Cache 一致性通信特别频繁,故保留 Mesh 结构, 而在 Cache 层则采用长连线的方式.如图 5(b)所示,采用传统三维 Mesh 结构从 A 节点传送消息 到 B 节点需要 7 拍,而如图 5(c)所示,若采用长连 线,则从节点 A 到节点 B′再到节点 B 只需要 3 跳, 而该拓扑结构是由整数线性规划(ILP)根据目 标函数和约束函数计算得出.由于该方法只适合 36 个节点以下的网络,作者还提出了两个扩展规 模的方法.这种设计降低了网络直径且具有较高 的性能,但是不规则的长连线设计也增加了布局 复杂度.



图 5 低阶低直径片上网络结构

前面介绍的这几种拓扑逻辑都是在三维对称结构的基础之上针对三维集成电路的特点设计优化的,而下面这3种拓扑逻辑均是由二维片上网络直接引入到三维片上网络中,同时能够非常契合三维结构的特点,具有各自的优势.

在二维片上网络中,树形结构是一种常用的拓扑结构,根据应用需求,可分为蝶形胖树、一般胖树和遗传胖树等^[36,50-51].Feero等人^[22]将树形结构引入到三维片上网络.这种胖树结构通过层次化的连接具有较小的网络直径,并且能达到非常高的性能提升.然而这种设计存在的缺点是后端设计较为复杂且温度控制不够理想.

De-Brujin 图是一种直径小、路由简单且具有容错等特性的网络拓扑,为了获得简单、可靠、吞吐率高与延时短等优势,该拓扑首先被应用在二维片上网络^[52].电子科技大学的 Chen 等人^[25]设计了一种在水平面和垂直平面上均采用 De-Brujin 图作为拓扑结构的三维片上网络.在大多数情况下,基于 De-Brujin 图的三维片上网络结构能够同时取得延时和功耗的最佳性能,可以解决由于网络规模增大导致延迟增大的问题,具有良好的可扩展性.该方法的缺点是性能效率不高,传统的一些布局布线算法无法复用,且对节点分布也有要求.

二维蜂巢结构^[53]相比于传统的矩形 Mesh 或 Torus 结构,可降低大约 40%的网络开销. 芬兰图 尔库大学的 Yin 等人^[26]根据蜂巢片上网络的这一 特点,提出了图 6 的三维蜂巢片上网络结构,同时设 计了一种死锁避免路由算法.该设计实际上是从二 维片上网络延伸到三维片上网络.增加了维度后,在 水平方向上依然保持蜂巢拓扑结构;在垂直方向上, 为控制路由节点的阶,作者将节点分为奇数节点和 偶数节点两类,一类只有向上的连线,一类只有向下 的连线,通过该方法进一步降低硬件开销和功耗.然 而这种结构也无法采用传统的布局布线算法,同时 平均节点跳数较高,影响性能.



上述拓扑结构都各具特点,基于 Mesh 的三维 片上网络形状规则,布局布线相对来说难度不大,非 常适合于商业化产品的应用,但是这一类拓扑结构又 均存在网络延时大、功耗开销大等问题.其他基于非 Mesh结构的三维片上网络,如低直径低阶网络、基 于 De-Brujin 图的网络结构虽然性能高、直径小,但 又存在芯片布局布线复杂、实现难度较大等问题.因 此,未来三维片上网络拓扑结构设计将面向可扩展性 强、布局布线简单、充分利用垂直空间等方向发展.

3.2 路由算法

路由算法定义了一个消息从源节点到目的节点的 路径,分为确定性路由、健忘路由和自适应路由3类, 根据传输路径的长度也被划分为最短路由或非最 短路由.路由算法的性能对于网络吞吐率和延迟有 着直接影响,越好的路由算法越能均衡网络负载. 对于三维片上网络的路由算法,要求其满足简 单、高性能和鲁棒性三个关键目标.同时,随着三维 片上网络设计规模的不断增大,系统出现错误的概 率也相应提高,需要更多的容错设计来解决这些问 题以提高可靠性.此外,三维集成电路对于异构体系 结构的支持,使得片上网络通信流量形成一定的特 征,如何通过路由算法均衡网络负载,提高带宽利用 率并降低功耗也成为路由算法研究的一大热点.现 有的路由算法研究主要集中在这两方面,根据路由 算法的分类分别介绍容错路由、全局流量均衡、层间 流量均衡和低功耗路由4个路由算法,其路由思路 和功能特点见表1.

表 1 三维路由算法

路由算法	路由思路	功能特点
AFRA 容错路由算法 ^[31]	$Z-XY \rightarrow X-Z-XY$	对 TSV 容错且可靠性高
流量平衡路由算法[32]	根据网络状态自适应	缓解拥塞并降低热效应
层复用健忘路由算法[33]	Z-XY-Z/Z-YX-Z	层间负载均衡且结构简单
面向功耗免死锁路由算法[34]	分8个象限,功耗自适应	功耗低且无死锁

AFRA 是一种低开销的三维 Mesh 片上网络容 错路由算法^[31],主要针对竖直 TSV 出现的永久性 错误而设计.该路由算法在默认情况下采用 Z-XY 算法,若 Z-XY 路径的 Z 方向出现了永久性错误,则 采用 X-Z-XY 路由算法.第1步,消息将先从源节点 传递到 X 方向最短路径上距离源节点最近的逃脱 节点上,第2步,从该逃脱节点再经过 ZXY 路由送 到目的节点.若在最短路径上没有合适的逃脱节点, 则会选择最小 ID 号的节点作为逃脱节点.在死锁避 免机制上,该文证明了当所有 Z 方向的链路错误是 同向的,则无需虚通道避免死锁.该算法是对 Z-XY 算法的改进,开销较低且可靠性较高,并具有较好的 性能.该算法的局限性在于只能支持竖直方向的容 错,且全是同向错误时才无需虚通道.

为解决三维片上网络的拥塞问题,缓解热效应,

Chang 等人^[32]提出了一种流量平衡的路由策略.该路由策略包括一个高级临时转接路由算法,该算法 主要采用 3 个方法:(1)收集网络中的拥塞和节流 信息以均衡负载;(2)减少路径复杂度以降低平均 延迟;(3)调整路由路径来散热.这种动态的路径选 择算法简单高效,不仅具有较低的延迟,同时有利 于网络的流量均衡以及热管理.但是在随机流量 模型下,部分节点具有非常高的通信频率,性能并 不好.

当三维片上网络的各层负载不均衡时,效率非常低.因此,Ramanujam等人^[33]提出了一种多层复用 三维片上网络结构,这种结构其实是一种 Z-XY-Z 或 Z-YX-Z 的健忘路由算法,可有效实现各层的负 载均衡.该文献用一种输入分层与输出层恢复的结 构替代传统的一层一跳结构.如图 7 所示,报文从节



图 7 层复用网络结构示意图(Rk 为第 k 层路由器对应的 IP 模块)

点注入后,首先进入输入分层阶段,该输入分层结构 连接对应于该位置的各层的路由器,根据路由器报 文计数器状态,将报文随机分配到负载较轻的一层, 通过维序路由到达目的节点对应的该层节点,再进 入输出层恢复结构,由交叉开关分配到目标节点所 在层的路由器.虽然多层复用三维片上网络结构具 有负载均衡和路由器结构简单等优点,但该结构仍 存在两个缺点,一是在竖直方向上需要两跳通过,二 是增加的输入分层和输出层恢复结构是两个4×4 的交叉开关,每次报文通信需要多经过两个交叉开 关,降低了性能.

在控制网络功耗并避免死锁方面,南京大学的 虞潇等人^[34]提出一种面向功耗免死锁三维全动态 路由算法.该算法将二维平面上的奇偶转弯模型扩 展到三维结构,将路由空间划分为8个象限,在每个 象限选取对应的路由策略,即先走不同的轴,到达对 应的象限后采用自适应路由,一方面避免互相之间 的打扰;另一方面也避免回路的产生从而避免死锁. 该算法还通过比较三维空间中6个方向的功耗指标 选择节点功耗较小的路径从而降低功耗,并使报文 分布均匀.与传统的 Z-XY 算法、静态 XYZ 算法和 TTAR^[54]算法相比,该设计在功耗方面显著下降,但 是在平均延迟和网络吞吐率等性能方面却略有损失.

三维片上网络的路由算法已经不仅是为了将消息从源节点传输到目的节点,它还能发挥容错、负载 均衡等功能性的作用.三维芯片密度增加,热效应问 题严重,对于芯片可靠性的要求更高,同时三维片上 网络相对于二维片上网络在路径多样性上有了很大 提升,有更多的空间和路径来实现容错、负载均衡的 功能.因此,未来三维片上网络路由算法的研究应面 向多功能应用算法研究,提高可靠性和降低功耗将 与提高性能一样重要.

3.3 路由器微结构

传统的路由器微结构主要包括如下模块:注入 缓存、路由计算模块、虚通道分配器、交叉开关分配 器、交叉开关以及输出单元等.路由器的功能主要是 通过流水线来提高吞吐率.在片上网络中每一个路 由器延迟占了总的通信延迟的很大一部分,需要通 过减少路由器流水线段来降低延迟并提高吞吐率.

由于三维片上网络的规模增大,维数增加,无论 是路由器端口数量还是交换开关的复杂度都不可避 免地增加,因此,三维片上网络的路由器微结构研究 重点是降低交叉开关复杂度等方面.图 8显示了以 降低开销为目的的三维片上网络路由器微结构设计 方向.现有的研究一类通过三维分解降低路由器交 叉开关的开销;一类通过减少端口的数量来降低路 由器开销;另一类则通过层次化设计实现三维高阶 Switch.



图 8 三维片上网络路由器微结构

文献[20]考虑一种纯三维交叉开关的实现,即 将交叉开关三维化.如图9所示,采用6个晶体管将 交叉开关的连接点向三维扩展,这种连接点称为连 接盒,垂直链路嵌入到交叉开关中,不需要额外的物 理链路,性能非常好.例如一个报文从底层的东输入 端口传播到顶层的南输出端口仅需要一拍.尽管性 能非常高,连通性非常好,但这种设计致命的缺点是 在交叉开关的连接点增加了大量垂直链路,造成了 严重的路径冗余和复杂度开销.对于一个位宽为1 的完整三维交叉开关,需要25个连接盒,由于每个连 接盒需要 6 个晶体管,则一个交叉开关需要 150 个 晶体管. 控制这些晶体管又需要至少 150 个控制信 号,一个4×4×4的三维片上网络结构,则需要接近 10000个晶体管.同时,其路径复杂度又严重过剩, 在真实应用程序中,这些连接的利用率非常低.因 此,对于延迟和面积约束非常严格的片上网络路由 器来说,这种设计是不现实的.





DimDe 路由器微结构^[20]是对纯三维交叉开关的改进.由于纯三维交叉开关开销过大,考虑将其化简成三个模块,即将 XYZ 三个维度分解成行模块、列模块和纵模块,分别仲裁.如图 10 所示,该设计将5×5 的交叉开关简化为行模块和列模块的两个 4×2的交叉开关,尽管增加了交叉开关的数量,但其规模依然小于一个单体的 5×5 交叉开关.在仲裁设计上,对于一个 XYZ 维序路由算法,第1 步先仲裁单层本地报文,需要传播到其他层的报文胜出,第2 步

东入

西入

开销.

上/下

行模块

纵模块

上/下

图 10 DimDe 交叉开关

将这些纵向传播的所有报文进行全局仲裁,该仲裁

器的设计是为了满足更多的并行通信. DimDe 三维

片上网络结构相对于总线结构提供了较高的全局

带宽,相对于纯三维片上网络结构降低了大量硬件

▶ 东出

▶ 西出



列模块

北入

南入

低路由器开销的目的,其结构如图 11 所示.该结构 基于 LastZ 路由规则,即非层内的报文传播,最后一 跳必须是竖直方向上的.因此,作者根据这个规则做 了两点结构设计:(1)去掉一个输入端口,将 6×6 的路由器简化为 5×6 的路由器结构;(2)为每个处 理核单元设计一个外包结构,该结构由一个有限状 态机和一个两路选择器组成.通过这样的结构设计, 使得报文在最后一跳直接从 dTDMA 总线跳到处 理核,而无需经过最后一个路由器.这种设计相对于 前面介绍的三维片上网络和总线混合结构,具有较 小的面积开销,平均节点跳数降低了一跳,简化的路 由器结构使得路由器可以工作在更高的频率.虽然 一定程度上缓解了总线拥塞问题,但总线无法并行 通信依然是这种结构的瓶颈,导致在一些特定流量 模型中效率降低.





在三维高阶路由器的研究方面,密歇根大学 的 Blaauw 领导的研究小组提出了一种支持单周 期仲裁的高阶三维集成的 Hi-Rise Switch^[44]. 该高 阶 Switch 基于 2D Swizzle-Switch^[55] 和三维折叠 Switch 结构^[56]设计,一方面将仲裁逻辑集成到交叉 开关的交叉点中,使得仲裁器复用数据通路,降低开 销;另一方面优化三维折叠 Switch 结构,如图 12 所 示将其划分为层次化的 Switch,本地 Switch 负责每 层的报文分配,层间 Switch 负责层间的报文分配. 在设计上关注两个点,一个是降低垂直通路的数量, 控制 TSV 数量并降低开销;另一个是层次化结构带 来的分配不公平问题,设计了一种基于类的最近最 少分配仲裁策略(CLRG),为每个输入配置一个三优 先级的计数器来实现. 该 Switch 支持 64 阶 4 层结构 的实现,相对于二维设计和传统三维折叠 Switch 有 更高的性能和更低的开销. Hi-Rise Switch设计在



图 12 Hi-Rise Switch 结构图 (N为端口数目,L为层数、C为任意两层间链路数目)

Mesh 拓扑结构中能实现支持千核的超大规模三维结构.该 Switch 的不足是垂直链路根据输入或输出分档,每个输入都固定到一条垂直通道,当多个输入竞争该通道时容易拥塞,该策略不够公平,仍有提升空间.

现有的三维片上网络路由器微结构研究以降低 硬件开销为主要目的,这是由于三维结构相对于二 维结构性能已经有了很大提升,而三维结构面临的 硬件复杂度、热效应等问题必须以降低开销作为主 要解决方式.因此,未来短期内路由器微结构的研究 将依然集中于实现低复杂度低功耗的路由器结构.

3.4 2.5 维硅中介技术

硅中介(Silicon interposer)技术即 2.5 维堆叠

技术是在三维堆叠技术尚未成熟前的一种过渡技术,相对来说能较好地支持现有的一些设计工具和 设计方法,目前也已成为三维片上网络研究者重点 关注的研究方向.如图 13 所示,2.5 维集成电路是 将多个同构或异构的部件相邻地堆叠在硅中介层 上,相邻部件之间通过硅中介层进行通信.通过 2.5 维堆叠可以大量增加存储容量,同时避免三维堆叠 中的热效应.现在已经出现了许多商用 2.5 维堆叠 产品^[57].硅中介技术的研究目前依然是比较新的方 向,其主要研究如何利用硅中介层进行通信,硅中介 层作为多个异构部件层的下一层,负责这些部件间 的通信.



图 13 2.5 维堆叠集成电路

如图 14(a)所示,在传统的 2.5 维堆叠设计中, 只会利用一小部分硅中介层的连线资源,作为相邻 部件间的通信.多伦多大学的 Jerger 等人^[35]开发利 用了硅中介层的大量未用资源,达到性能的大幅提 升.该文献提出了一种如图 14(b)所示的通用方法 来利用硅中介层的连线资源,使得片内通信或片间 通信均可以利用硅中介层的连线资源.在实现细节 上一方面区分访存报文和一致性消息报文,使之分 别利用不同层的网络进行通信,另一方面通过负载 均衡方法和快速通路合理控制并均衡了网络流量, 同时还降低网络直径.该方法在带来非常小的额外 资源开销的同时大幅度提高了性能,可以有效缓解 网络拥塞,同时大幅提高了等分带宽.作者还提出了



一些拓扑逻辑在各层的设计方法以及报文转发 策略.

该文献第一次开发了硅中介层的设计空间,许 多现有的技术都能在此应用,以达到较高的性能,具 有非常广阔的前景.2.5 维片上网络硅中介层的下 一步研究一方面可以从结构和算法设计入手,对硅 中介层拥塞控制、容错、路由算法等进行深入研究; 另一方面,可以将 2.5 维片上网络硅中介技术应用 到更多的环境,例如 CPU、GPU、其他硬件加速器以 及 DRAM 等组合的环境里,更好地实现异构集成, 将其潜能更大地挖掘出来.

3.5 热管理

热效应不仅是三维集成电路,同样也是三维片 上网络所面临的严峻挑战.在二维集成电路或是二 维片上网络中热效应并不明显,但当更多地设备器 件层竖直堆叠在一起时,功耗密度和热传导路径的 长度均增加.高温所带来的问题是更长的传播延迟 和更高的静态功耗.在超出芯片工作范围的温度工 作可能会导致错误输出和可靠性降低等问题.此外, 热效应使得散热和封装成本均增加.目前对于热效 应的研究也非常多,在系统结构层的研究一方面通 过结构设计降低功耗和热的产生,另一方面通过路 由控制热点区域报文流向消除热点,平衡通信量.

"国立"台湾大学的 Huang 等人^[41]提出了一种 共享路由器的三维片上网络的热管理结构设计,通 讨降低路由器数量来降低芯片温度,解决热效应问 题. 由于远离散热器的 m 层路由器即使是在空闲状 态仍然会产生较大的静态功耗,并导致热效应的问 题. 作者将远离散热器的 m 层存储和 PE 模块的路 由器去掉,通过 TSV 总线将这些部件与其他 N-m 层的路由器相连,实现路由器共享.该策略与传统三 维结构相比,在80℃的温度约束下可将报文接收率 提高两倍,并节省13.4%的功耗开销.

在热效应严重的三维芯片中,温度每升高10℃, 连线延迟将增加大约5%,而漏电流将呈指数增长. 宾夕法尼亚州立大学的 Addo-Quave^[39]提出采用遗 传算法来进行热和通信感知的三维片上网络应用任 务的映射和布局,相比干随机映射布局方法,该方法 热管理效果较好,明显降低了系统通信量.因此,功 耗和温度的增加都得到直接缓解,对三维片上网络 性能可靠性有着直接影响.

高温给系统性能、功耗、可靠性和开销等带来了 一些不可避免的问题,针对三维片上网络的热效应 这一问题, Chao 等人在文献 [40] 提出了一个三维片 上网络流量和热感知的实时热管理策略.该策略包 括两个部分:(1) 主动向下路由策略:(2) 垂直反馈 节流策略.主动向下路由策略是一种流量感知策略, 将部分流量迁移到底层,靠近散热器的同时,分散拥 塞流量,垂直反馈节流策略则是一种热感知策略,将 过热的路由器节流以提升网络可靠性,但该策略对 三维片上网络的性能有一定的影响.

热管理的研究在二维片上网络研究中就被提 出,由于堆叠散热困难,在三维片上网络中热管理的 重要性凸现出来.对热管理的研究一方面可以缓解 热效应的问题;另一方面也能提高系统可靠性.目前 对于三维片上网络热管理的研究主要在系统层,通 过全局策略进行控制,未来热管理的研究还将更多 地面向微结构层和电路层,从底层缓解和消除热效 应问题.

3.6 TSV 设计优化

三维堆叠垂直互连的几种技术中,TSV 最有潜 力提供最大的连接密度. TSV 连接技术并不限制芯 片堆叠的数量,但堆叠后的热效应和产能是制约其 发展的关键因素.如图 15 所示,TSV 连接技术可以 用在 Bulk 工艺或 SOI 工艺下, SOI 相对于 Bulk 工 艺具有相对更小的垂直中心距(Vertical Pitch). TSV 技术的优势非常多,然而也存在一些缺点,包 括面积开销大、TSV 产能较低、寄生电容问题、路由 拥塞问题等,目前在三维片上网络垂直互连研究领 域,对于 TSV 连接技术的研究和发掘最为活跃,主 要目的是提高 TSV 可靠性,减少引脚数量和面积以 控制通孔周围的应力问题.



图 15 TSV T.艺

对于竖直方向的 TSV,竖直连接的引脚面积对 于垂直互连的开销影响很大,文献[46]提出了一种 竖直 TSV 连接的串行通信方法以减少 TSV 数量. 该串行通信方法基于移位寄存器,由一个串行发射 器和一个串行接收器组成.数据通过一条串行传播 通路进行传播.作者测试了不同的串行程度对性能 的影响,在4:1或2:1的串行程度下较为实际,即 采用4条或2条串行传播通路传播数据.在这种串 行程度下性能损失和功耗开销都较小,引脚面积可 大幅减少.该设计使得竖直方向的热 TSV 分布更为 合理,峰值温度也相应降低,节省的空间方便了芯片 核的布局规划.但这种设计的缺点是需要增加逻辑 复杂度完成串行通信,且对性能有所影响.

文献[47]也通过减少垂直 TSV 连接的引脚数 量,来减少面积和热功耗,该文献作者提出了一种三 维片上网络结构 BBVC,即一种基于双向双同步垂 直通道的三维片上网络.该结构可以动态的自配置 竖直链路的方向来传递报文,在竖直方向用一个高 频率双向通道替代一对单向通道以达到更高的带宽 利用率、更小的引脚面积和更高的可靠性.图 16 为 BBVC 结构的示意图,为垂直端口配置控制模块,使 之具有双向通信能力.BBVC 结构的优点是显著减 少了 TSV 引脚面积,简化了布局,同时异构集成能 力强;缺点则是增加了竖直方向上的跳步数,路由器 交叉开关开销也较大.





文献[48]为了减少 TSV 高开销这一缺陷,设计 了一种性能与开销较合算的三维片上网络结构.该 结构的目的就是为了减少 TSV 的个数.通过对每一 层的节点进行分类,将节点归到几个虚拟组,每个虚 拟组连接一个 TSV 通信通道,并设计了一系列算法 来管理通信,减少垂直 TSV 出现空闲的时间,以提 高 TSV 的带宽利用率,继而提高全局性能.该方法 可以显著减少 TSV 面积且性能下降非常小.

国内对于 TSV 设计优化的研究也取得了显著 的成果.合肥工业大学的欧阳一鸣等人^[49]提出了一 种 TSV 复用容错策略.传统的容错方法是增加一定 数目的冗余链路,然而针对 TSV 故障,由于 TSV 是 数量的限制,该方法难以适用.因此,作者提出通过 对故障链路中功能良好的 TSV 复用,并通过将数据 微片多次传输的方式实现对 TSV 硬故障的容错.此 外还通过添加 ECC 编码解码模块来实现瞬时故障 容错功能.该策略一定程度地控制了硬件开销,时延 和吞吐率都有提高,但是由于容错策略复用 TSV, 出现错误时,微片在垂直方向分多次传输,导致零负 载延迟增加.

上述 TSV 的设计优化研究均是以降低 TSV 开 销或控制 TSV 数量提高可靠性为目的,这是由于目 前的制造工艺使得 TSV 的产品良率较低,成本较 高.然而在现有技术条件下,垂直方向上采用 TSV 互连性能较高且最容易商业化应用,所以目前受到 广泛关注.未来对于垂直 TSV 互连的体系结构研究 一方面继续对其设计优化降低开销;另一方面还需 要挖掘更多更有前景的垂直互连技术,如无触点或 交流耦合互连通信等.

3.7 其他研究

除了前面所介绍的6个研究热点外,对于三维 片上网络的研究还包括新型芯片堆叠方式、设计 方法学研究、任务映射、缓存布局、三维光互连研 究^[58-59]以及性能评估等许多方面.由于篇幅限制,本 文无法全面完整地介绍,只针对具有代表性的新型 芯片堆叠方式、设计方法学研究和缓存布局三个方 面的研究进行介绍.

Puttaswamy和Loh在文献[37]中提出一种三 维处理器设计,其每个独立功能模块贯穿多个芯片 层,虽然这一结构在当前技术条件下还存在一定困 难,但随着多层堆叠处理器技术的成熟,这种结构非 常灵活有效.MIRA 三维片上网络结构是由美国宾 夕法尼亚大学的Park等人^[38]基于该技术提出的一 种多层路由器结构.在逻辑结构上,其与二维 Mesh 结构的路由器类似,在竖直方向上,其路由器结构贯 穿多层网络,与二维 Mesh 相比具有较小的节点面 积和路由器间较短的距离.此外.路由器无需增加额 外的功能结构,只需要将功能模块合理地布局在多 层结构中.这种结构具有非常高的性能,但是其针对 的是三维设计处理器核,难以重用目前已经高度优 化的二维处理器核设计技术,且 Switch 设计复杂度 也很高.

在三维片上系统中,如何设计一个能效高的片 上网络结构并满足三维堆叠技术的诸多约束是研究 者在体系结构设计时最为关心的问题之一.为了解 决这个问题,文献[36]提出了一种高能效三维片上 网络设计的综合方法.该方法根据核的位置以及三 维片上网络的各种约束,决定最合适的拓扑结构和 各层路由器的位置.三维片上网络的诸多约束除技 术约束外,还包括针对特定应用的网络通信特征.这 种定制的三维片上网络拓扑结构针对其通信特征可 以大幅度的减少功耗开销并降低传输延迟.

在三维多核片上系统中,一般将共享存储部件 设置在顶层.当 DRAM 都在三维片上网络的顶层 时,一般都采用 Wide IO 接口进行通信^[60],外部存 储端口不再是片上系统的性能瓶颈.新的性能瓶颈 将是网络热点导致的链路拥塞和饱和.文献[60]针 对这一新瓶颈,提出了一个新的性能评估分析方 法^[61]来分配链路缓存,以这种方式缓解热点区域的 拥塞.该方法在网络设计阶段,通过采用 Uniform Random 和 Hotspot 混合流量模型寻找网络中可能 存在的热点区域,计算并分配各链路所需的合适的 缓存数量,以提升网络性能.然而,这种方法不具有 实时控制能力,链路缓存分配在设计阶段已确定,因 此不能根据网络状态进行调整.

本节主要从系统层、微结构层和电路层 3 个层 次出发,介绍了三维片上网络体系结构的最新研究 进展.事实上,这些研究热点之间并不是相互独立 的,它们的关系错综复杂,每一个研究点都涉及其他 多个研究点,而几乎所有的三维片上网络研究最终 都回归到拓扑结构、路由算法以及路由器微结构设 计等问题.TSV的设计优化、低功耗和热管理、提高 可靠性和性能、2.5 维硅中介技术、容错以及负载均 衡等方面的研究,都是通过针对拓扑结构、路由策略 和路由器微结构做具体设计来实现的;同时所有针 对拓扑结构、路由策略和路由器微结构的设计,其目 的又都是为了提高性能和可靠性、实现容错、降低功 耗等.不同于二维片上网络研究,三维片上网络的研 究结合其自身结构的特点以及面临的问题,更多地 出现在热管理、TSV 设计优化、2.5 维硅中介层利 用以及对层间极短距离的利用上.这些研究的成果 能够更好地发挥三维结构的优势特点.表 2 则将本 文介绍的文献以这种方式呈现出来.

	拓扑结构	路由算法	路由器微结构
提高性能	[21][22][23] [24][25][35] [36]	[32][33][35]	[38]
降低开销	[36][53]	[33]	[20][42][43] [44][48]
容错和提高 可靠性		[31][40]	[47][49]
均衡负载		[32][33][35]	
热管理和 低功耗		[23][32][34] [40]	[41][46]
TSV 设计优化	[21]		[46][47] [48][49]
硅中介利用	[35]		
缓存分配			[60]

表 2 三维片上网络体系结构研究分类

4 研究方法与产品研发

4.1 模拟环境与性能测试

在片上网络体系结构研究中,其巨大的设计空间需要模拟器来有效地评估其性能,并选择合适的参数优化设计.二维片上网络和三维片上网络的主要功能都是连接各节点实现通信,三维片上网络体系结构模拟器目前主要沿用二维结构的模拟器.

由于在体系结构研究中,我们更多地关心链路 层以上层次的结构,只需要考虑节点连通性、链路延 迟、链路带宽以及路由器内部结构等设计,而较少地 关心 TSV 等结构的具体实现.因此,将二维片上网 络模拟器升级为三维片上网络模拟器,最主要是在 拓扑结构的增维.需要在片上网络垂直方向增加相 应的链路,同时在路由器微结构增加相应的端口.垂 直方向的链路主要模拟 TSV 的设计,其延时相应较短,垂直链路的带宽位数代表了 TSV 的数量,因此还需要注意垂直链路数量、密度和带宽的控制.而路由算法、流控机制、网络流量模型等方面则可根据设计相应实现.

现有的片上网络体系结构模拟器按照实现编程 语言划分为三类,包括以C++实现的Booksim^[62]、 Garnet^[63]、OPENC-Sim^[64];以SystemC实现的Noxim^①、 Nirgam^[65]以及以verilog 或VHDL等硬件描述语 言实现的模拟器.这些模拟器可以模拟片上网络的 拓扑结构、路由算法、路由器微结构、流控机制、网络 流量模型等各种机制和特性,通过对代码的修改,可 以实现各种设计,并测试其性能和硬件开销.此外, 这些片上网络模拟器还可以链接到Gem5等全系 统模拟器或Trace模拟器Netrace^[66]、SynFull^[67]等 实现全系统环境下测试片上网络的性能.除了上 述性能模拟器之外,还有专门测试片上网络功耗 和面积的模拟器,如Orion^[68]、DSENT^[69]和RTL Router^[70]等.

三维片上网络专用模拟器 VNOC3² 是由 Cristinel Ababei 领导的小组采用 C++语言开发的 仿真平台.该仿真平台是研究两层或三层的三维片 上网络架构的一个框架,建立在一个 VNOC 仿真平 台和一个 B* 树布局之上.该平台是专门针对三维片 上网络结构设计的,然而在灵活性和功能全面性方 面不及 Booksim 模拟器.

此外,"国立"台湾大学的 Access IC Lab 也开发 了一个三维片上网络联合模拟平台 Access Noxim³.其代码结构集成了 Noxim和 Hotspot两 个模拟器,因此该平台集成了网络模型、功耗模型和 热模型.该模拟平台采用 SystemC 语言实现,对于 重视热管理的设计实现有很高的应用价值.

4.2 产品研发

由于三维集成电路依然面临着众多挑战,对于 三维集成电路的研发还处于先期发展阶段,三维逻 辑堆叠芯片的商业化产品并没有问世.目前的三维 堆叠产品主要是对于存储部件的堆叠以及采用 2.5 维硅中介层技术的堆叠.

在实验研究型芯片方面,如表 3 所示,2007 年

① Noxim: Network-on-chip simulator, 2008. http://sourceforge.net/projects/noxim

[@] VNOC3,2009. http://venus.ece.ndsu.nodak.edu/ $\sim cris/software.html$

③ Access Noxim, 2015. http://access.ee.ntu.edu.tw/noxim/ index.html

Intel 推出了实验性三维芯片 Teraflops 处理器; 2008 年罗切斯特大学的 Friedman 领导的小组研发 了第一款堆叠芯片,主频达到 1.4 GHz;乔治亚理工 学院研制的 3D-MAPS 三维芯片^[71],包含了一层逻 辑层和一层 DRAM 层,采用 130 nm 工艺,主频仅 277 MHz.其逻辑层 64 核的核间通信采用的是 2D Mesh 结构,垂直互连方面采用了 Tezzaron 的 1.2 µm 直径的 TSV 技术. 密歇根大学研制的 Centip3De 是 一款大规模三维 CMP 处理器^[72],采用两层结构,一 层为包含 64 个 ARM Cortex-M3 核心的系统,该 64 核被组织成 16 个群,通过 crossbar 相连接,另一 层是 Cache. 两层通过 TSV 总线连接. IBM 的 Wordeman 等人^[73]研制了一个三维原型系统,由一 层存储层和一层逻辑层组成. 其中存储结构采用的 是 eDram,其存储逻辑来自 IBM Power7 processor L3 Cache,其垂直互连采用的是三维 Switch 结构.

表 3 研究型堆叠芯片

产品/技术	出现时间	单位	堆叠类型	堆叠层数
Teraflops 处理器	2007	Intel	实验性存储堆叠	_
Rochester Cube	2008	罗切斯特大学	逻辑堆叠	_
3D-MAPS	2011	乔治亚理工学院	存储+逻辑堆叠	$2 \sim 5$
Centip3De	2012	密歇根大学	存储+逻辑堆叠	2
3D System Prototype	2012	IBM	存储+逻辑堆叠	2

目前,三维逻辑堆叠的芯片由于技术原因并没有 产品化,主要是作为实验性芯片由一些研究机构进行 研究开发.由于堆叠层数都非常少,一般都是两层堆 叠,故其片上网络结构相对简单,多为在逻辑层的 2D Mesh 或 Switch,垂直互连采用总线或 Switch 为主.

商业化产品方面,如表4所示,目前最主要的两种存储堆叠标准分别是由美光公司领导设计的HMC(Hybrid Memory Cube)^①和AMD和Hynix领导设计的HBM(High Bandwidth Memory)^②.它们均是采用TSV互连堆叠的高性能RAM接口标准,应用到各种芯片的实现上.HBM与HMC的区别体现在DRAM运行频率、总位宽、发热以及扩展性层面.在应用方面,采用HBM的AMDFijiXT已

经于 2015 年 6 月问世. 英伟达 2016 年将推出的下 一代 GPU 架构的 Pascal 也将采用 HBM. HMC 则 将广泛地应用到 Intel 第二代至强 Phi 处理器以及 三星存储堆叠芯片等产品中. 在国内的产品研发方 面,武汉新芯在 2015 年 5 月宣布其 9 层结构的三维 存储器芯片通过了基本功能的电学验证,预计 2017 年将实现量产. 在 2.5 维堆叠集成方面,赛灵思公司 则在 2011 年推出了第一款 2.5 维堆叠的 Virtex-7 2000T FPGA^[57]. 该 FPGA 芯片在硅中介层上装配 了 4 块 FPGA 子芯片,通过硅中介层互连打破了 I/O 瓶颈,提供了 FPGA 间 10 000 条数据路径的通信. 实现同样的性能,比起二维结构面积只有四分之一, 功耗更是只有六分之一.

表 4 商业堆叠芯片

产品/技术	出现时间	单位	堆叠类型	堆叠层数	量产时间
HMC	2011	美光	存储堆叠	$4 \sim 8$	2014
HBM	2009	AMD/Hynix	存储堆叠	$4 \sim 8$	2015
三维数据型闪存	2015	武汉新芯	存储堆叠	9	2017
3D V-NAND	2014	三星	存储堆叠	32	2014
3D NAND Flash	2015	东芝	存储堆叠	48	2016
Virtex-7	2011	赛灵思	2.5 维堆叠	2	2011

商业化产品主要出现在存储堆叠上,由于其结构较为规整,热效应并不严重,因此技术逐渐成熟, 并出现了许多商业产品.在其互连结构方面,以 HMC标准为例,底层存储控制器通过高阶 Switch 相连,上层存储部件通过总线,与对应的存储控制器 连接,组成三维存储结构.

5 挑战与展望

目前,三维片上网络依然面临着许多问题和挑

战,除了本文第二部分提到的三维片上网络的三大 问题之外,其进一步发展还面临一系列的挑战和技 术瓶颈^[74-75],主要包括:

(1)缺乏商业化的 EDA 工具^[76]和设计方法.由 二维结构的设计转换到三维结构的设计,片上网络 结构从空间上获得了非常多的结构优势和性能优 势,同时,在设计上所要考虑的约束条件和设计因素

① HMC Specification 1.0, 2013. http://www.hybridmemorycube.org

② High Bandwidth Memory(HBM), 2015. http://www.amd. com/zh-cn/innovations/software-technologies/hbm

也不断增加. 然而目前没有成熟和合适的 EDA 工 具和设计方法支持三维片上网络设计,这直接导致 三维堆叠的商业化产品无法实现. 随着设计空间的 增大,设计空间探索难度急剧增加^[77],如何发挥三 维堆叠优势,设计高性能的三维堆叠产品也困难重 重,因此,对于支持三维集成电路的 EDA 工具和设 计方法的研究势在必行.

(2) 散热问题. 芯片堆叠所产生的热效应来源 于功耗密度的急剧增加,它带来的散热问题目前是 限制三维集成电路及其片上网络发展的最大瓶颈, 对于芯片可靠性和稳定性的影响非常大,许多设计 都受到散热问题的限制而被迫进行缩减,三维片上 网络也因此需要各种结构设计来降低开销. 解决该 问题需要研究者从电路层到系统层提出思路,对三 维堆叠的热效应建立模型,研究可能的解决办法以 迁移热效应.

(3) 通孔周围的应力问题. 在三维片上网络的 垂直 TSV 周围会出现应力问题,限制了 TSV 阵列 与器件的间距以及 TSV 在竖直方向布局密度,从而 影响网络带宽和器件布局. 这就要求在垂直 TSV 的 设计中考虑全局性因素,需要与电源、时钟网络等因 素共同考虑网络资源分配,同时需要在三维片上网 络进行体系结构设计与优化.

(4)测试问题.目前研究者对于三维堆叠技术 的测试问题的研究还非常不足,缺乏针对三维片上 网络的测试方法.测试问题最大的局限就是对于堆 叠的硅片难以探测访问,难以判断堆叠芯片内部网 络的情况,为高可靠性产品的实现带来了隐患.因 此,三维堆叠及三维片上网络技术急需研究者提出 测试设计的解决办法.

上述挑战虽然带来了许多困难,但同时也指引 了三维片上网络发展的方向,现在和未来的研究除 了提高三维片上网络的性能之外,主要研究目标就 是解决这些问题.结合三维片上网络研究热点以及 系统结构设计的未来发展,三维片上网络体系结构 层设计主要有两个大的发展方向.一方面是将二维 片上网络已有的结构或机制引入或移植到三维片上 网络中,并根据三维结构特点进行优化设计,包括拓 扑结构、路由算法、路由器微结构等方面,这对于三 维片上网络研究的快速发展进步是非常有利也非常 有必要的;另一方面,在三维片上网络的设计中最重 要的一点就是注重混合设计.由于片上网络的三维 结构在水平方向上的两维与竖直方向的第三维并不 是对称的,因此在设计时可以分别结合片内和片间 网络的特点进行不对称设计,发挥不同结构各自的 优势特点,这种不对称性既可以适用于拓扑结构、路 由器微结构等物理结构上,也适用于路由算法、流控 机制等通信流量的管理上,从而更好地发挥出三维 片上网络结构的优势.

综上所述,三维片上网络由于具备二维片上网络所不具备的空间优势,其发展空间非常广阔,必将伴随三维集成电路的发展成为下一代多核甚至众核系统的主流片上通信架构.

参考文献

- [1] Kumar H, Singh S, Kumar C. 3D integration-opportunities and challenges. International Journal of Electronic and Electrical Engineering, 2010, 3(3): 125-131
- [2] Wulf W A, McKee S A. Hitting the memory wall: Implications of the obvious. ACM SIGARCH Computer Architecture News, 1995, 23(1): 20-24
- [3] Black B, Annavaram M, Brekelbaum N, et al. Die stacking (3D) microarchitecture//Proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture. Orlando, USA, 2006: 469-479
- [4] Ramm P, Klumpp A, Weber J. 3D integration technologies for MEMS/IC systems//Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting. Capri, Italy, 2009: 138-141
- [5] Topol A W, La Tulipe D C, Shi L, et al. Three-dimensional integrated circuits. IBM Journal of Research and Development, 2006, 50(4.5): 491-506
- [6] Xia Yan. Present situation and development of 3D integration. China Integrated Circuit, 2011, 20(7): 23-28(in Chinese)
 (夏艳. 3D 集成的发展现状与趋势. 中国集成电路, 2011, 20(7): 23-28)
- [7] Loh G H, Xie Y, Black B. Processor design in 3D die-stacking technologies. IEEE Micro, 2007, 27(3): 31-48
- [8] Loi G L, Agrawal B, Srivastava N, et al. A thermally-aware performance analysis of vertically integrated (3-D) processormemory hierarchy//Proceedings of the 43rd Annual Design Automation Conference. San Francisco, USA, 2006; 991-996
- [9] Xie J. Zhao J. Dong X, et al. Architectural benefits and design challenges for three-dimensional integrated circuits// Proceedings of the 2010 IEEE Asia Pacific Conference on Circuits and Systems. Kuala Lumpur, Malaysia, 2010: 540-543
- [10] Xie Y, Cong J, Sapatnekar S S. Three-Dimensional Integrated Circuit Design. New York: Springer, 2010
- [11] Davis W R, Wilson J, Mick S, et al. Demystifying 3D ICs: The pros and cons of going vertical. Design & Test of Computers, 2005, 22(6): 498-510
- [12] Xie Y. Processor architecture design using 3D integration technology//Proceedings of the 23rd International Conference on VLSI Design. Bangalore, India, 2010: 446-451

- [13] Dally W J, Towles B P. Principles and Practices of Interconnection Networks. Elsevier, Morgan Kaufmann, USA, 2004
- [14] Hemani A, Jantsch A, Kumar S, et al. Network on chip: An architecture for billion transistor era//Proceedings of the IEEE NorChip Conference. Turku, Finland, 2000, 31: 1-8
- [15] Li Li, Zhang Yu-Ang, Fu Yu-Xiang, et al. The study of memory architectures for 3D chip multi-processors. Journal of Nanjing University(Natural Sciences), 2014, 50(3): 330-335(in Chinese)

(李丽,张宇昂,傅玉祥等.三维众核片上处理器存储架构研究.南京大学学报(自然科学),2014,50(3):330-335)

- [16] Liu C C, Chen J H, Manohar R, et al. Mapping system-onchip designs from 2-D to 3-D ICs//Proceedings of the IEEE International Symposium on Circuits and Systems. Kobe Japan, 2005; 2939-2942
- [17] Hu J, Marculescu R. Energy-and performance-aware mapping for regular NoC architectures. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2005, 24(4): 551-562
- [18] Jeang Y L, Wey T, Wang H Y, et al. Mesh-tree architecture for network-on-chip design//Proceedings of the 2nd International Conference on Innovative Computing, Information and Control. Kumamoto Japan, 2007: 262
- [19] Deng Y, Maly W P. Interconnect characteristics of 2.5-D system integration scheme//Proceedings of the 2001 International Symposium on Physical Design. Sonoma County, USA, 2001: 171-175
- [20] Kim J, Nicopoulos C, Park D, et al. A novel dimensionallydecomposed router for on-chip communication in 3D architectures//Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA'07). New York, USA, 2007: 138-149
- [21] Li F, Nicopoulos C, Richardson T, et al. Design and management of 3D chip multiprocessors using network-inmemory//Proceedings of the 33rd Annual International Symposium on Computer Architecture. New York, USA, 2006, 34(2): 130-141
- [22] Feero B S, Pande P P. Networks-on-chip in a three-dimensional environment: A performance evaluation. IEEE Transactions on Computers, 2009, 58(1): 32-45
- [23] Matsutani H, Koibuchi M, Amano H. Tightly-coupled multi-layer topologies for 3-D NoCs//Proceedings of the IEEE International Conference on Parallel Processing. Xi'an, China, 2007: 75
- [24] Xu Y, Du Y, Zhao B, et al. A low-radix and low-diameter 3D interconnection network design//Proceedings of the IEEE 15th International Symposium on High Performance Computer Architecture. Raleigh, USA, 2009; 30-42
- [25] Chen Y, Hu J, Ling X. De Bruijn graph based 3D Network on Chip architecture design//Proceedings of the International Conference on Communications, Circuits and Systems. Milpitas, USA, 2009: 986-990
- [26] Yin A W, Xu T C, Liljeberg P, et al. Explorations of honeycomb topologies for network-on-chip//Proceedings of the 6th IFIP International Conference on Network and Parallel Computing. Gold Coast, Australia, 2009: 73-79

- [27] Tan Hai, He Yue-Shun, Jin Wen-Bing, et al. V-Mesh: A low-latency and energy-efficient network-on-chip structure for 3D stacked many-core. Chinese Journal of Computers, 2014, 37(10): 2139-2152(in Chinese)
 (谭海,何月顺,靳文兵等. V-Mesh: 面向三维堆叠芯片的低时延低功耗片上网络结构. 计算机学报, 2014, 37(10): 2139-2152)
- [28] Zhang Da-Kun, Song Guo-Zhi, Wang Lian-Lian, et al. Survey on topologies of three-dimensional network on chip. Journal of Frontiers of Computer Science & Technology, 2015, 9(2): 129-164(in Chinese)
 (张大坤,宋国治,王莲莲等. 三维片上网络拓扑结构研究综 述. 计算机科学与探索, 2015, 9(2): 129-164)
- [29] Rahmani A M, Latif K, Liljeberg P, et al. Research and practices on 3D networks-on-chip architectures//Proceedings of the NORCHIP. Tampere, Finland, 2010: 1-6
- [30] Chen Yi-Ou, Hu Jian-Hao, Ling Xiang. Research on topologic architecture of three-dimensional network on chip. Telecommunications Science, 2009, 25(4): 39-44(in Chinese) (陈亦欧,胡剑浩,凌翔. 三维片上网络拓扑研究. 电信科 学, 2009, 25(4): 39-44)
- [31] Akbari S, Shafiee A, Fathy M, et al. AFRA: A low cost high performance reliable routing for 3D mesh NoCs// Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE). Dresden, Germany, 2012: 332-337
- [32] Chang W C, Tseng H W, Kuo C F. A traffic-balanced routing scheme for heat balance in 3D networks-on-chip// Proceedings of the 29th Annual ACM Symposium on Applied Computing. New York, USA, 2014: 1437-1442
- [33] Ramanujam R S, Lin B. A layer-multiplexed 3D on-chip network architecture. IEEE Embedded Systems Letters, 2009, 1(2): 50-55
- [34] Yu Xiao, Li Li, Zhang Yu-Ang, et al. A power-aware dead lock avoid three-dimensional full-adaptive routing algorithm for 3D NoC. Acta Electronica Sinica, 2013, 41(2): 329-334 (in Chinese)
 (虞潇,李丽,张宇昂等. 一种面向功耗免死锁三维全动态 3D NoC 路由算法. 电子学报, 2013, 41(2): 329-334)
- [35] Jerger N E, Kannan A, Li Z, et al. NoC architectures for silicon interposer systems: Why pay for more wires when you can get them (from your interposer) for free?//Proceedings of the 47th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO). Cambridge, England, 2014: 458-470
- [36] Murali S, Seiculescu C, Benini L, et al. Synthesis of networks on chips for 3D systems on chips//Proceedings of the 2009 Asia and South Pacific Design Automation Conference. Piscataway, USA, 2009: 242-247
- [37] Puttaswamy K, Loh G H. Thermal herding: Microarchitecture techniques for controlling hotspots in high-performance 3D-integrated processors//Proceedings of the IEEE 13th International Symposium on High Performance Computer Architecture. Scottsdale, USA, 2007: 193-204
- [38] Park D, Eachempati S, Das R, et al. MIRA: A multilayered on-chip interconnect router architecture//Proceedings

of the 35th Annual International Symposium on Computer Architecture. New York, USA, 2008: 251-261

- [39] Addo-Quaye C. Thermal-aware mapping and placement for 3-D NoC designs//Proceedings of the International SOC Conference. Herndon, USA, 2005: 25-28
- [40] Chao C H, Jheng K Y, Wang H Y, et al. Traffic-and thermal-aware run-time thermal management scheme for 3D NoC systems//Proceedings of the 2010 4th ACM/IEEE International Symposium on Networks-on-Chip (NOCS). Grenoble, France, 2010: 223-230
- [41] Huang Y R, Pan J H, Lu Y C. Thermal-aware routersharing architecture for 3D Network-on-Chip designs// Proceedings of the 2010 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS). Paris, France, 2010: 1087-1090
- [42] Rahmani A M, Liljeberg P, Plosila J, et al. LastZ: An ultra optimized 3D networks-on-chip architecture//Proceedings of the 14th Euromicro Conference on Digital System Design (DSD). Oulu, Finland, 2011: 173-180
- [43] Rahmani A, Liljeberg P, Plosila J, et al. Partial-LastZ: An optimized hybridization technique for 3D NoC architecture enabling adaptive inter-layer communication//Proceedings of the 2012 International IEEE Conference SoC Design Conference (ISOCC). Jeju Island, Korea, 2012: 281-284
- [44] Jeloka S, Das R, Dreslinski R G, et al. Hi-Rise: A highradix switch for 3D integration with single-cycle arbitration// Proceedings of the 47th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO). Cambridge, England, 2014: 471-483
- [45] Matthias T, Burggraf J, Burgstaller D, et al. Thin die stacking for wide I/O memory-on-logic. Solid State Technology, 2012, 55(5): 17-20
- [46] Pasricha S. Exploring serial vertical interconnects for 3D ICs//Proceedings of the 46th Annual Design Automation Conference. San Francisco, USA, 2009; 581-586
- [47] Rahmani A M, Liljeberg P, Plosila J, et al. BBVC-3D-NoC: An efficient 3D NoC architecture using bidirectional bisynchronous vertical channels//Proceedings of the 2010 IEEE Computer Society Annual Symposium on VLSI (ISVLSI). Lixouri Kefalonia, Greece, 2010: 452-453
- [48] Hwang Y J, Lee J H, Han T H. 3D Network-on-Chip system communication using minimum number of TSVs// Proceedings of the 2011 International Conference on ICT Convergence (ICTC). Seoul, Korea, 2011: 517-522
- [49] Ouyang Yi-Ming, Yang Yi-Ze, Liang Hua-Guo, Huang Zheng-Feng. Reuse and fault-tolerant strategies for TSVs of three-dimensional network-on-chip. Journal of Electronic Measurement and Instrument, 2013, 27(3): 229-235(in Chinese)

(欧阳一鸣,杨懿泽,梁华国,黄正峰.三维片上网络 TSV 复用容错策.电子测量与仪器学报,2013,27(3):229-235)

- [50] Greenberg R I, Guan L. An improved analytical model for wormhole routed networks with application to butterfly fat-trees//Proceedings of the 1997 International Conference on Parallel Processing. Bloomington, USA, 1997: 44-48
- [51] Grecu C, Pande P P, Ivanov A, et al. A scalable communicationcentric SoC interconnect architecture//Proceedings of the 5th

International Symposium on Quality Electronic Design. San Jose, USA, 2004: 343-348

- [52] Hosseinabady M, Kakoee M R, Mathew J, et al. Reliable network-on-chip based on generalized de Bruijn graph// Proceedings of the IEEE International Conference on High Level Design Validation and Test Workshop. Irvine, USA, 2007: 3-10
- [53] Stojmenovic I. Honeycomb networks: Topological properties and communication algorithms. IEEE Transactions on Parallel and Distributed Systems, 1997, 8(10): 1036-1042
- [54] Lin Shu-Yen, Yin Tzu-Chu, Wang Hao-Yu, et al. Trafficand thermal-aware routing for throttled three-dimensional network-on-chip systems//Proceedings of the 2011 International Symposium on VLSI Design, Automation and Test. Taiwan, China, 2011: 320-323
- [55] Satpathy S, Sewell K, Manville T, et al. A 4.5 Tb/s 3.4 Tb/s/W 64×64 switch fabric with self-updating leastrecently-granted priority and quality-of-service arbitration in 45nm CMOS//Proceedings of the 47th Annual IEEE/ACM International Symposium on Solid-State Circuits Conference Digest of Technical Papers (ISSCC). San Francisco, USA, 2012: 478-480
- [56] Sewell K, Dreslinski R G, Manville T, et al. Swizzle-switch networks for many-core systems. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2012, 2(2): 278-294
- [57] Saban K. Xilinx stacked silicon interconnect technology delivers breakthrough FPGA capacity, bandwidth, and power efficiency. Xilinx White paper: Vertex-7 FPGAs, 2011: 1-8
- [58] Morris R, Kodi A K, Louri A. Dynamic reconfiguration of 3D photonic networks-on-chip for maximizing performance and improving fault tolerance//Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture. Washington, USA, 2012; 282-293
- [59] Ramini L, Bertozzi D. Power efficiency of wavelength-routed optical NoC topologies for global connectivity of 3D multi-core processors//Proceedings of the 5th International Workshop on Network on Chip Architectures. New York, USA, 2012: 25-30
- [60] Foroutan S, Sheibanyrad A, Pétrot F. Cost-efficient buffer sizing in shared-memory 3D-MPSoCs using wide I/O interfaces// Proceedings of the 49th Annual Design Automation Conference. New York, USA, 2012: 366-375
- [61] Foroutan S, Thonnart Y, Hersemeule R, et al. An analytical method for evaluating network-on-chip performance// Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE). Dresden, Germany, 2010: 1629-1632
- [62] Jiang N, Becker D U, Michelogiannakis G, et al. A detailed and flexible cycle-accurate network-on-chip simulator// Proceedings of the 2013 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). Austin, USA, 2013; 86-96
- [63] Kumar A, Agarwal N, Peh L S, et al. A system-level perspective for efficient NoC design//Proceedings of the IEEE

International Symposium on Parallel and Distributed Processing. Miami, USA, 2008: 1-5

- [64] Jueping C, Gang H, Shaoli W, et al. OPNEC-Sim: An efficient simulation tool for Network-on-Chip communication and energy performance analysis//Proceedings of the 2010 10th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT). Shanghai, China, 2010: 1892-1894
- [65] Jain L, Al-Hashimi B M, Gaur M S, et al. NIRGAM: A simulator for NoC interconnect routing and application modeling//Proceedings of the Design, Automation and Test in Europe Conference. San Diego, USA, 2007: 16-20
- [66] Hestness J, Grot B, Keckler S W. Netrace: Dependencydriven trace-based network-on-chip simulation//Proceedings of the 3rd International Workshop on Network on Chip Architectures. New York, USA, 2010: 31-36
- [67] Badr M, Jerger N E. SynFull: Synthetic traffic models capturing cache coherent behaviour//Proceedings of the 2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA). Minneapolis, USA, 2014: 109-120
- [68] Kahng A B, Li B, Peh L S, et al. ORION 2.0: A fast and accurate NoC power and area model for early-stage design space exploration//Proceedings of the Conference on Design, Automation and Test in Europe. European Design and Automation Association. Leuven, Belgium, 2009; 423-428
- [69] Sun C, Chen C H O, Kurian G, et al. DSENT A tool connecting emerging photonics with electronics for optoelectronic networks-on-chip modeling//Proceedings of the 2012 6th IEEE/ACM International Symposium on Networks on Chip (NoCS). Copenhagen, Danmark, 2012: 201-210
- [70] Becker D U. Efficient Microarchitecture for Network-on-



LI Chen, born in 1989, Ph.D. candidate. His main research interest is 3D Networks-on-Chip.

Background

Networks-on-Chip (NoCs) have been proposed to be used in complex SoCs for communication between cores for more than ten years, and some commercial processors already deploy such networks. As semiconductor manufacturing processes are approaching the physical limits, it brings the community to explore three-dimensional (3D) VLSI design. The emerging 3D stacking technology has provided a new horizon for NoCs designs which brings tremendous advantages, including shorter global interconnects lengths, larger bandwidth, better scalability and heterogeneous integration. At the same time, there are still many challenges for 3D NoC, such as thermal issue, vertical connection issue, high router Chip Routers [Ph. D. dissertation]. Stanford University, Stanford, USA, 2012

- [71] Lim S K. 3D-MAPS: 3D massively parallel processor with stacked memory//Design for High Performance, Low Power, and Reliable 3D Integrated Circuits. Springer, New York, USA, 2013: 537-560
- [72] Fick D, Dreslinski R G, Giridhar B, et al. Centip3De: A 3930DMIPS/W configurable near-threshold 3D stacked system with 64 ARM Cortex-M3 cores//Proceedings of the 2012 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC). San Francisco, USA, 2012: 190-192
- [73] Wordeman M, Silberman J, Maier G, et al. A 3D system prototype of an eDRAM cache stacked over processor-like logic using through-silicon vias//Proceedings of the IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC). San Francisco, USA, 2012, 186-187
- [74] Bernstein K, Andry P, Cann J, et al. Interconnects in the third dimension: Design challenges for 3D ICs//Proceedings of the 44th Annual Design Automation Conference. New York, USA, 2007: 562-567
- [75] Tu K N. Reliability challenges in 3D IC packaging technology. Microelectronics Reliability, 2011, 51(3): 517-523
- [76] Tsai Y F, Xie Y, Vijaykrishnan N, et al. Three-dimensional cache design exploration using 3DCacti//Proceedings of the 2005 IEEE International Conference on Computer Design: VLSI in Computers and Processors. San Jose, USA, 2005: 519-524
- [77] Xie Y, Loh G H, Black B, et al. Design space exploration for 3D architectures. ACM Journal on Emerging Technologies in Computing Systems, 2006, 2(2): 65-103

MA Sheng, born in 1986, Ph. D., assistant researcher. His main research interest is computer architecture.

WANG Lu, born in 1991, M. S. candidate. Her main research interest is computer architecture.

GUO Yang, born in 1971, Ph. D., researcher, Ph. D. supervisor. His main research interests include microprocessor design & verification.

design complexity, etc.

Extending NoCs to the third dimension opens a new area of research. Hence, there is an increasing demand to classify and analyze the existing 3D NoC technologies. This paper first reviews advantages in 3D stacking technologies, and then introduces the current research progress in 3D NoC from system layer, microarchitecture layer and circuit layer. Finally, some open issues and future directions in 3D NoC are discussed.

This research is supported by the National Natural Science Foundation of China (61133007, 61303065) and the Doctoral Fund of Ministry of Education (20134307120028).